VIP- assignment 3
Hold 2 - Assignment group 5

Marcus Frehr Friis-Hansen: lns611
Mikkel Keller Andreasen: trq281
Søren Strøyberg: lds147

December 2025

1 Questions and answers

Write for a generic pixel, light source and normal, Lambert's law in its simplest form

This is given by the following formula from p.34. David A. Forsyth, 2012

$$I = \rho I_0 \cos\theta$$

Where I is the brightness of diffuse patch illuminated by a distant light source, projected onto the pixel. I_0 is the intensity of the light source, θ is the angle between the light source direction and the surface normal, ρ is the diffuse albedo, a value describing the amount of light reflecting from a diffuse surface.

How is Lambert's law modified to deal with self shadows.

$$I = \rho \max(s \cdot n, 0)$$

Lauze, n.d.

I works by using two vectors, s a vector pointing towards the light-source, and n a vector pointing perpendicular to the surface normal. The dot-product between these two vectors results in the angle relative to the two vectors. In the case of the surface pointing away from the light source, it clamps the angle to zero, as to avoid a negative lighting value.

$$I = \rho \max(s \cdot n, 0) + A$$

Lambert's law uses a distant light-source, implicit thought of as the sun, which to simplify is assumed to be infinitely far away, such that the rays cast by the *luminaire* (the light source), are parallel. This also means that surfaces that are shaded receive no light at all, which inst the case under real conditions. They usually receive light from other sources, such as light reflected from nearby surfaces. To make up for it, you could add a constant term to lambert´s law, such that shadows don't appear entirely dark. This model inst that successful though.

What about cast shadows? Comment on the difference between the two.

Self shadows are surfaces where the normal points away from the lights origin, whereas cast shadows are occluded because something is blocking the lighting source.

Comment on the modelling limits of Lambert's law.

This modified version of Lambert's law only works for self-shadows and not cast shadows, as lamberts law is local, and cast shadows aren't a local event. It can only be used on diffused surfaces i.e. surfaces where light scatters in all directions, such that the viewing angle doesn't matter. As stated earlier it also assumes a single light source casting parallel light rays onto the imaged surface, this is often not the case especially for instances where the imaged surface is illuminated artificially.

How can we obtain an estimate of albedo and normals in Woodham's approach to Photometric Stereo. Write the equation

The main idea is that we can isolate the albedo and surface normal from the brightness formula below, by using multiple pictures, taken using a fixed camera, but illuminated under different lighting conditions

$$I(u, v) = \rho(u, v) \cdot n(u, v) \cdot s$$

Using k images gives us a system of k equations, that we can solve using linear algebra. The matrix multiplication looks like so:

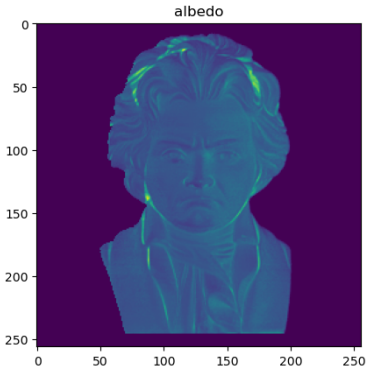
$$I = S \cdot g \qquad g = \rho n$$

We essentially just solve for g. If given 3 images $k = 3$ we can solve it exactly, and if $k > 3$ we can solve using the least squares method.

What should be done if one uses RANSAC. Please describe. It will help you when implementing the RANSAC based estimation

The idea of RANSAC(Random Sampling Consensus) is that if the data contains a lot of outliers, we can randomly sample to to find candidates for good estimates that we can use to produce a correct model, after many iterations.
When using RANSAC, I need to estimate the vector $m = \rho n$ in a way that is robust to outliers in the intensities. Based on the slides, this is done by repeatedly selecting three random measurements (the minimum needed to solve for m), computing a candidate solution, and then checking how many of the remaining measurements agree with this model within a threshold $|I_i - s_i^T m| < \tau$. I repeat this for a number of iterations and keep the model that gets the most inliers. In the end, I re-estimate m using only the inliers to obtain a stable and reliable result. Lauze, n.d.

2 Beethoven Dataset



(a) Albedo

(b) depth map

(c) depth map

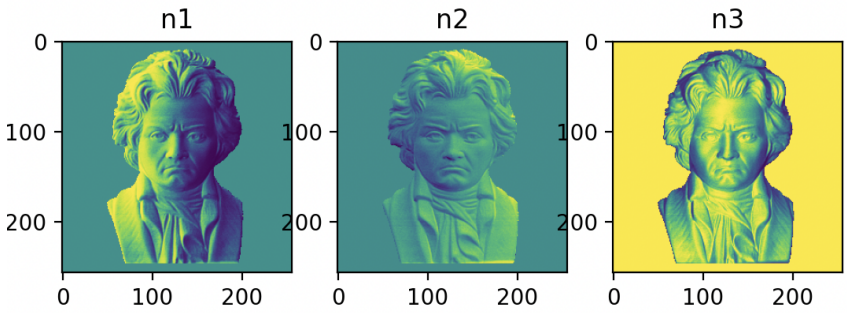


Figure 2: Normals

The Beethoven dataset contains three input images, a binary mask, and the 3×3 illumination matrix S. By using the mask, we collected all the non-zero pixels into a matrix $J \in \mathbb{R}^{3 \times n_z}$, where each column is the intensity values of a single pixel across the three images. The photometric stereo relation $J = SM$ was then inverted to obtain the albedo-modulated normals $M = S^{-1}J$.

We computed the albedo as the Euclidean norm of each column of M, and the normal field was extracted by normalizing M. The albedo image was reconstructed by inserting the albedo values back into an empty $(m \times n)$ image and visualized as shown above.

The depth map was obtained from the estimated normal field using unbiased surface integration. The reconstructed albedo, normal components, and integrated depth all show a consistent and stable reconstruction of the Beethoven bust, which is expected of a smooth Lambertian surface.

3 mat_vase Dataset

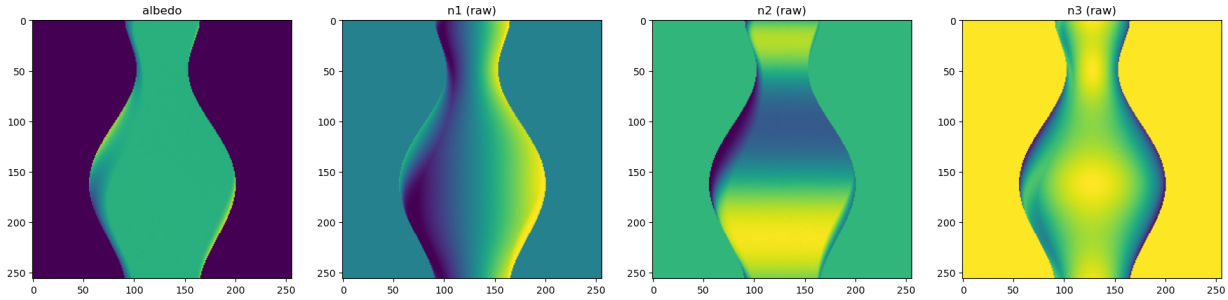


Figure 3: Reconstructed albedo and normal components for the mat_vase dataset. From left to right: albedo, n_1, n_2, and n_3.

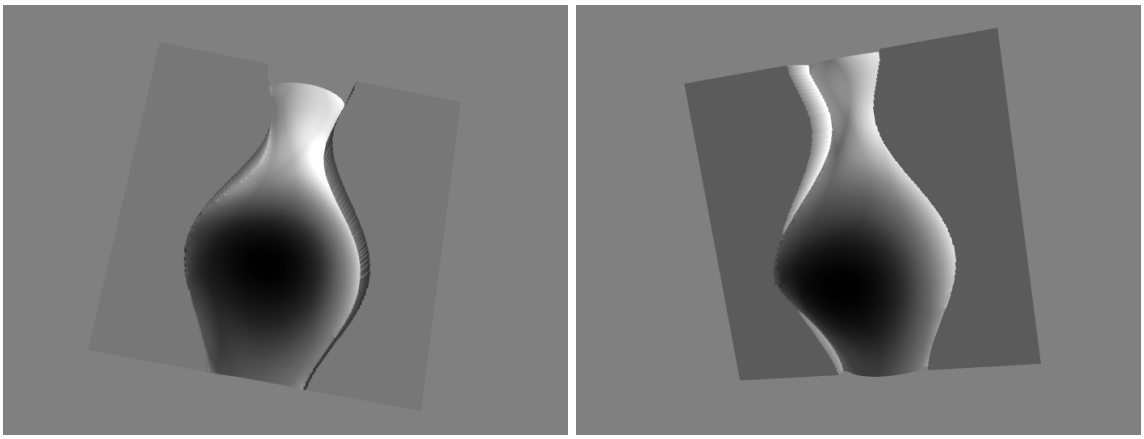


Figure 4: Depth map obtained by unbiased surface integration of the estimated normal field, visualized using Mayavi.

The mat_vase dataset is a clean and fully synthetic dataset consisting of three images, a binary mask, and a 3×3 illumination matrix S. As in the previous experiments, all valid pixels inside the mask were collected into an intensity matrix J, and the photometric stereo relation $J = SM$ was inverted to obtain the albedo-modulated normals $M = S^{-1}J$.

The albedo was computed as the norm of the columns of M, while the surface normals were obtained by normalization. The resulting albedo image is smooth and spatially consistent, which is expected for a matte, purely diffuse surface. The normal components also show clear and coherent structure across the vase.

The estimated normal field was integrated using unbiased surface integration. The resulting depth map, shown above, reveals a smooth and symmetric vase shape without noticeable artifacts, confirming a stable reconstruction for this dataset.

4 shiny_vase Dataset

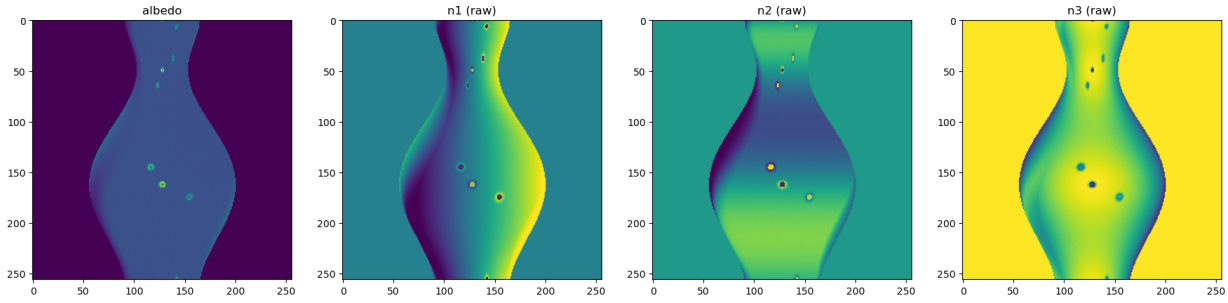


Figure 5: RANSAC-based reconstruction for the shiny_vase dataset. From left to right: albedo estimated with RANSAC, followed by the smoothed normal components n_1, n_2, and n_3.

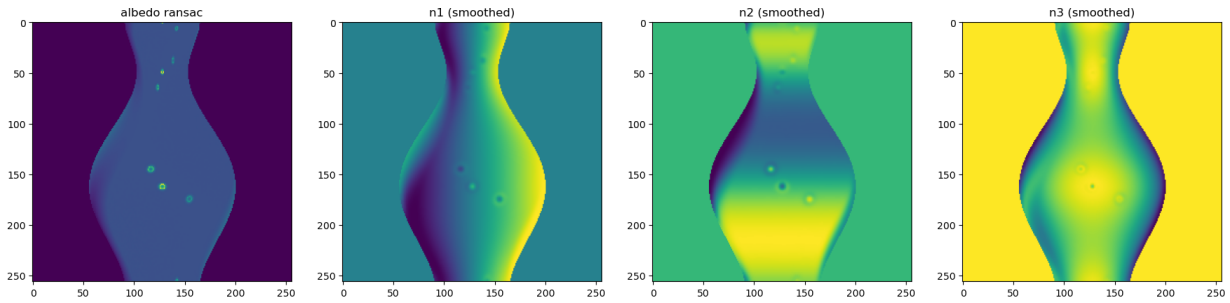


Figure 6: RANSAC-based reconstruction for the shiny_vase dataset. From left to right: albedo estimated with RANSAC, followed by the smoothed normal components n_1, n_2, and n_3.

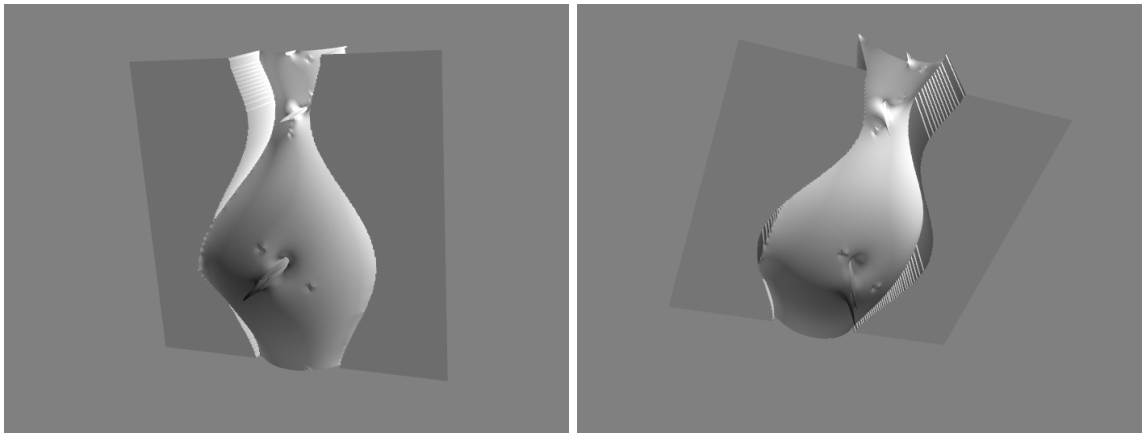


Figure 7: Depth map for the shiny_vase dataset visualized with Mayavi.

The shiny_vase dataset contains strong specular reflections that violate the Lambertian assumption and lead to unstable results when using standard photometric stereo. This shows up on the surface normals as spots. To handle this, we applied a RANSAC-based version of Woodham's method, which estimates the albedo-modulated normals by selecting intensity observations that best fit a model and discarding outliers. The resulting image looked almost identical to the those where we did not use RANSAC. It was only when we smoothed the image that these spots (artifacts) diminished. RANSAC therefore gives a better result when used together with smoothing, creating a visibly more coherent surface normal in all three directions.

The artifacts can also clearly be seen in the depth-map based on the photometric stereo without RANSAC, where they show up as sudden spikes, shooting off of the surface of the vase, which break the expected smooth shape. The artifacts can be seen on the images generated using RANSAC, but at a lesser degree when compared to the surface normals generated via Woodhams method. These are most obvious on the depth-map.

5 shiny_vase2 Dataset

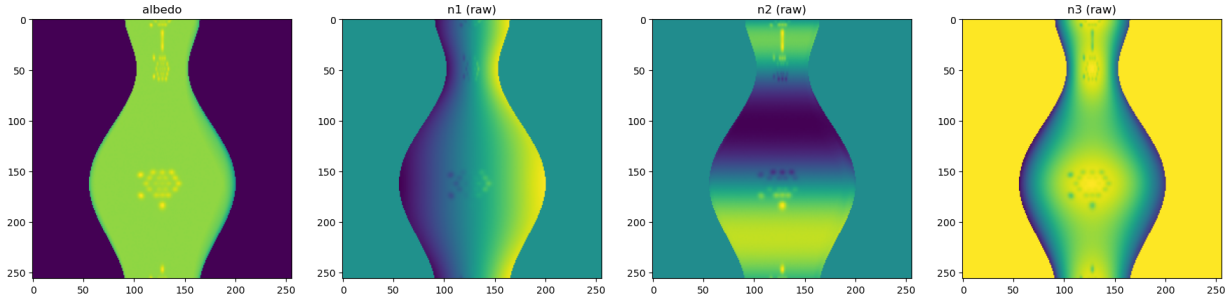


Figure 8: Standard photometric stereo reconstruction for the shiny_vase2 dataset. From left to right: albedo, n_1, n_2, and n_3 (raw).

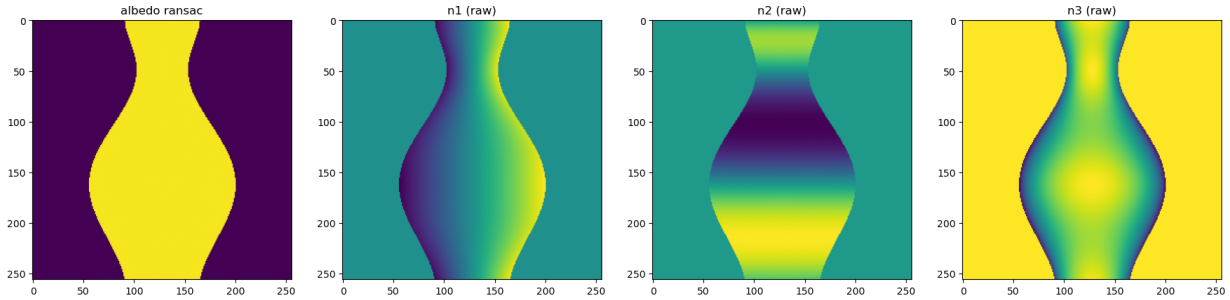


Figure 9: RANSAC-based reconstruction for shiny_vase2. From left to right: albedo estimated with RANSAC, followed by n_1, n_2, and n_3 (raw).

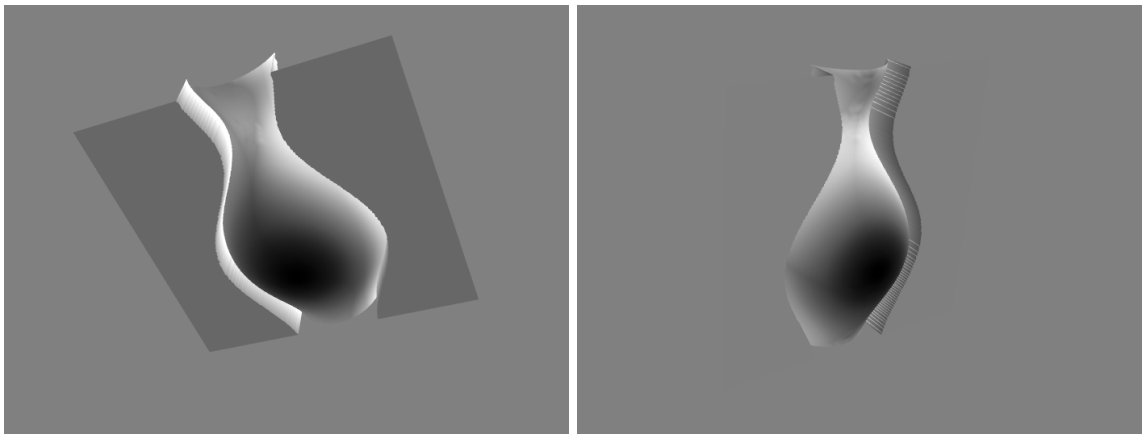


Figure 10: Depth map for the shiny_vase2 dataset, visualized with Mayavi.

The shiny_vase2 dataset uses the same setup as shiny_vase, but with a larger number of input images and lighting directions. The additional observations generally stabilize the normal estimation compared to the earlier shiny_vase case.

Despite this, the surface is still non-Lambertian, and specular highlights remain visible in the standard reconstruction, appearing as outliers in the albedo image and local artifacts in the raw normal components.

These artifacts are less significant when compared to the shiny_vase dataset. This is especially apparent when you look at the smoother and coherent depth map shown above. i.e. it has no visible peaks.

Applying the RANSAC-based method further reduces the influence of these highlights by rejecting inconsistent observations. As a result, the estimated albedo and normals become more consistent with no visible artifacts, it wasn't even necessary to perform any smoothing for the artifacts to disappear. This is also why we haven't displayed the smoothed results, as they provide no improvement in the surface normals. In this case RANSAC performed significantly better than Woodhams method.

6 Buddha Dataset

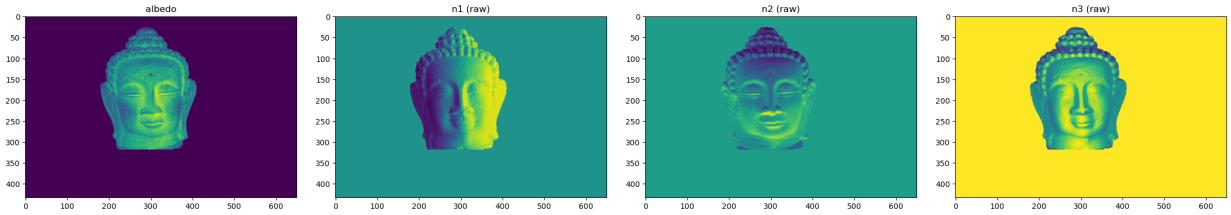


Figure 11: Standard photometric stereo reconstruction for the Buddha dataset. From left to right: albedo, n_1, n_2, and n_3 (raw).

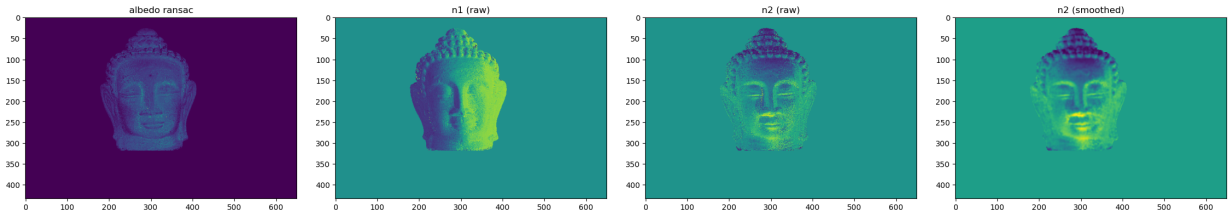


Figure 12: RANSAC-based reconstruction for the Buddha dataset. From left to right: albedo estimated with RANSAC, followed by n_1 (Raw), n_2 (Raw), and n_2 (smoothed).

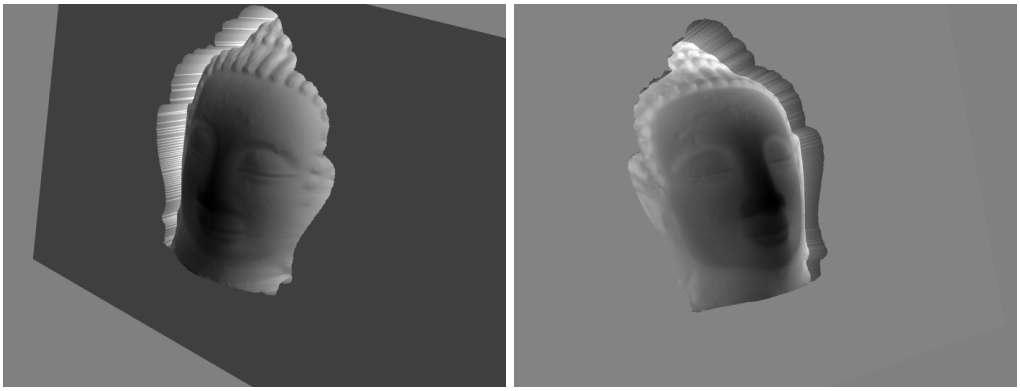


Figure 13: Depth map for the buddha dataset, visualized with Mayavi.

The Buddha dataset is more challenging than the synthetic vase datasets due to stronger 'grainy' surface texture, image noise, and relatively smaller dataset. In the standard photometric stereo reconstruction, this results in visibly noisy normal components. These effects are exacerbated when using RANSAC.
While Woodhams method tries to average out outliers, RANSAC tries to construct a homogeneous surface by

only using inliers. This results in a grainy surface texture visible in the surface normals. In the figure above showing the surface normals (n2) the surface texture degenerates to such a degree that it no longer looks coherent. This can be improved by smoothing, also visible in figure 11 above. The amount of smoothing iterations you perform on the surface normal has a big impact on the resulting image, as smoothing too much can make the surface normals look blurry.
The resulting depth map is constructed based on Woodhams method (not RANSAC), and looks pretty good. As the surface in the dataset inst specular, we don't see any extreme artifacts as those visible in the shiny_vase dataset, which is to be expected. We would say that Woodhams method performed better for the buddha dataset, which is expected for lambertian surfaces, while RANSAC performed poorly. This shows that RANSAC isn't ideal for noisy lambertian surfaces.

7 face Dataset

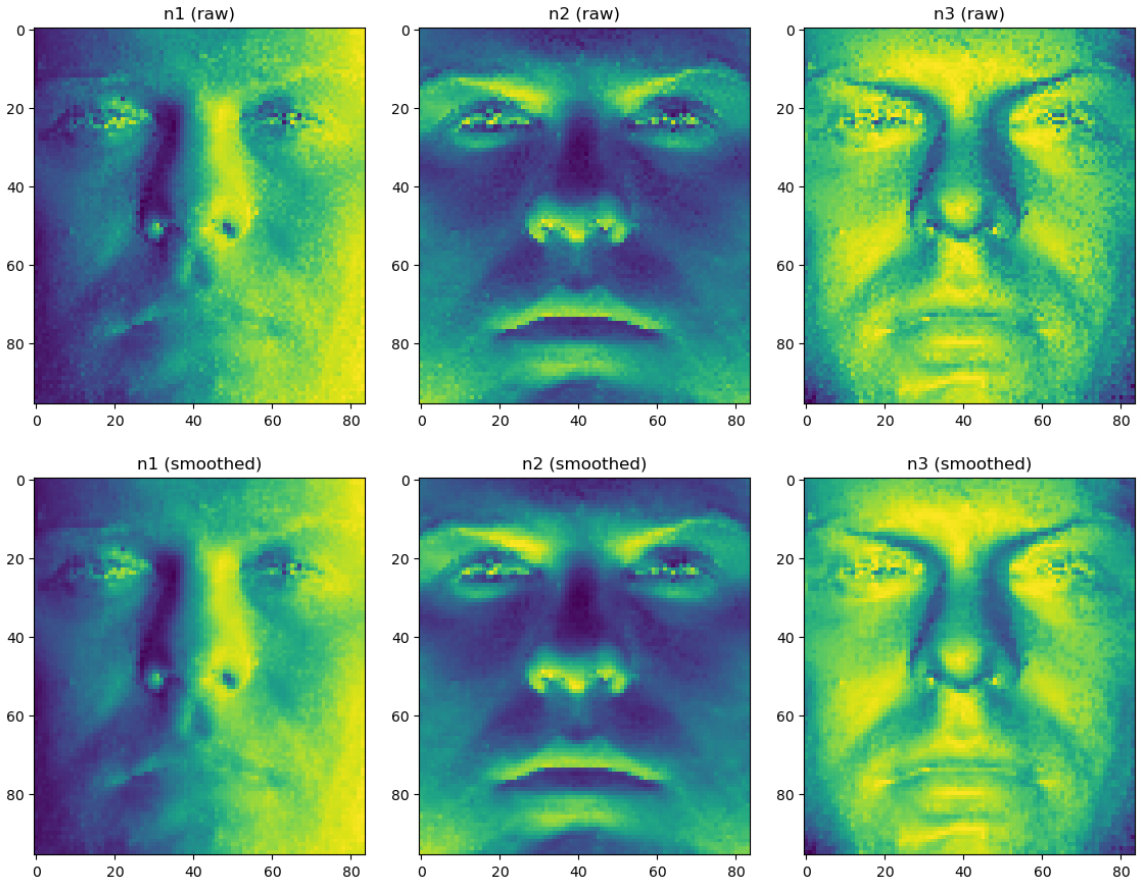


Figure 14: RANSAC-based normal estimation for the face dataset. Top row: raw normal components n_1, n_2, and n_3. Bottom row: corresponding smoothed normal components after applying normal field smoothing.

For the face dataset we directly applied the RANSAC based photometric stereo method. The surface contains some non-lambertian effects and fine scale texture, which makes Woodhams approach sensitive to outliers. RANSAC helps reduce the influence of such outliers by fitting the Lambertian model only to the most consistent observations for each pixel.

The raw normal components still contain noticeable high frequency noise, particularly around the eyes. To improve spatial consistency, we applied a smoothing step to the estimated normal field. As shown in

Figure 14, the smoothed normals preserve the overall facial structure while significantly reducing noise and small local artifacts.

Overall the combination of RANSAC and normal field smoothing produces a much more stable and visually coherent reconstruction for the face dataset, which underlines the importance of robust estimation and post processing when dealing with textured non-lambertian surfaces.

References

David A. Forsyth, J. P. (2012). *Computer vision a modern approach second edition*. Pearson.
Lauze, F. (n.d.). *Vision and image processing: Shading photometric stereo* [Slides].

-*